

Особые экономические зоны Российской Федерации: моделирование решений потенциальных резидентов и процесса их генерации

А. Е. Плесовских  

*Сибирский федеральный университет,
г. Красноярск, Россия*

 alexandermcme@gmail.com

Аннотация. В современных исследованиях широко обсуждается роль особых экономических зон в стимулировании экономического роста и развития России, формировании необходимых инвестиционных потоков и повышении инновационного потенциала страны за счет расширения производства продукции в высокотехнологичных отраслях экономики с высокой добавленной стоимостью. Цель исследования – моделирование процесса генерации резидентов и детерминация количественных факторов, оказывающих статистически значимый эффект на среднегодовой темп прироста компаний, участниц особых экономических зон в Российской Федерации. В работе описаны современные подходы в предсказании выбора потенциальных резидентов о начале ведения предпринимательской деятельности на территории ОЭЗ с применением классификационных подходов (Support Vector Machines, Decision Trees, Random Forest, Naive Bayes, K-Nearest Neighbor, Gradient Boosting) и регрессионных подходов (логистическая регрессия). Применен на практике современный классификационный алгоритм – Histogram-based Gradient Boosting Classification Tree, стабильный для анализа больших данных с отсутствующими значениями переменных, не требующий предварительной трансформации выборки. В работе подтверждена гипотеза о наличии положительной связи между местоположением организации и ее формируемым по итогам года финансовым результатом. В среднем по выборке компании-резиденты, расположенные вблизи центров субъектов РФ, более успешны по показателю генерируемой выручки. Гипотеза о наличии сильной взаимосвязи между показателями пространственной дифференциации субъектов РФ и показателями, характеризующими процесс генерации резидентов и частных инвестиций, не была в полной мере подтверждена. С практической точки зрения, результаты исследования могут быть применены как организациями-резидентами, потенциальными резидентами, так и управляющими компаниями ОЭЗ. Теоретическая значимость исследования заключается в спецификации предложенной модели бинарного выбора потенциальных резидентов, которая может быть расширена и обобщена в будущих работах. В настоящее время имеются все необходимые предпосылки для создания условий по развитию промышленности, высокотехнологичных отраслей экономики и выпуска продукции с высокой добавленной стоимостью с целью повышения устойчивости российской экономики.

Ключевые слова: особые экономические зоны России; процесс генерации резидентов; машинное обучение; регрессия и классификация; модели бинарного выбора.

1. Введение

В последние годы в России наметилась тенденция по созданию крупных экономических кластеров, направленных на стимулирование экономического роста и развития. Так, создание особых экономических зон (ОЭЗ) в Российской Федерации призвано выступить гарантом привлечения крупных инвесторов и создать условия для ведения успешной предпринимательской деятельности.

При формировании особых экономических зон приоритет отдается отраслям с высокой добавленной стоимостью, высокотехнологичным секторам экономики, служащим драйвером разработки технологий и новых видов продукции. Федеральный закон «Об особых экономических зонах в Российской Федерации» предусматривает такие типы создаваемых зон, как промышленно-производственные (ПП), технико-внедренческие (ТВ), туристско-рекреационные (ТР) и портовые (П) (табл. 1).

При формировании перечня потенциальных резидентов организации принимают решение о возможности или невозможности вступления в особую экономическую зону и начале деятельности на ее территории. Отмечается наличие у них преференциальных условий: 1) освобождение на определенный период от земельного, имущественного и транспортного налогов; 2) наличие льготных ставок по налогу на прибыль организаций; 3) действие таможенных льгот; 4) возможности пользования преимуществами, которые дает территория особой экономической зоны (возможность бесплатного подключения и льготные тарифы на электроэнергию, водо-, тепло- и газоснабжение; 5) развитые логистическая и транспортная сети (при наличии).

Все эти преференциальные условия увеличивают вероятность принятия положительного решения потенциальным резидентом. Вместе с тем потенциальные резиденты оценивают уровень необходимых инвестиционных вложений в течение

периода осуществления предпринимательской деятельности на территории особой экономической зоны, стоимость аренды необходимых промышленных площадей и некоторые другие параметры, влияющие на исход решения.

Ожидается, что фирмы, принимая решение о начале осуществления предпринимательской деятельности на территории особой экономической зоны, ориентируются на такие факторы, как величина инвестиций, которые необходимо будет осуществить в случае вступления в особую экономическую зону, различие величин налоговых ставок, которые отличаются для предприятий, являющихся резидентами ОЭЗ, и предприятий, не входящих в ОЭЗ, по определенным видам налогов. Кроме того, фирмы оценивают местоположение – текущее при осуществлении деятельности вне территории зоны и предполагаемое при вступлении в нее как фактор, оказывающий влияние на расположение потенциальных потребителей, на удаленность места работы потенциальных работников и на доступность транспортных магистралей, необходимых для доставки ресурсов и готовой продукции.

При оценке финансовых показателей менеджеры фирмы учитывают эффект от ее вступления на территорию особой экономической зоны на величину выручки, чистой прибыли, рентабельности продаж, активов и собственного капитала. Стоит также отметить, что некоторые показатели эффективности работы особых экономических зон, как было предположено в исследовании, находятся в тесной зависимости от величины социального капитала, различающегося в том или ином регионе страны.

Такие факторы вносят немаловажный вклад в успешность или неуспешность функционирования особых экономических зон. Различия в величине социального капитала выявляются в исследовании Ramos-Rodriguez et al. [1]. Мы исходим из предположения о том,

Таблица 1. Матрица сравнения типов ОЭЗ

Table 1. Comparison matrix of SEZ types

Линия сравнения	ПП-зоны	ТВ-зоны	ТР-зоны	П-зоны
Типы резидентов	Коммерческие организации, за исключением унитарных предприятий	ИП или коммерческие организации, за исключением унитарных предприятий	ИП или коммерческие организации, за исключением унитарных предприятий	Коммерческие организации, за исключением унитарных предприятий
Верхнее ограничение площади территории зоны	40 км ²	4 км ²	Без ограничения	50 км ²
Специфика экономической деятельности	Производство и переработка товаров и их реализация, а также оказание услуг по обеспечению перевозок и складированию товаров	Инновационная деятельность, создание, производство и реализация научно-технической продукции, создание и реализация программ для электронных вычислительных машин, баз данных, топологий интегральных микросхем, информационных систем, оказание услуг по внедрению и обслуживанию такой продукции	Строительство, реконструкция, эксплуатация объектов туристской индустрии, объектов, предназначенных для санаторно-курортного лечения, медицинской реабилитации и отдыха граждан	Портовая деятельность и строительство, реконструкция и эксплуатация объектов инфраструктуры морского порта, речного порта или аэропорта

Источник: составлено автором исследования на основании ФЗ «Об особых экономических зонах в Российской Федерации» от 22.07.2005 № 116-ФЗ¹.

что переменная, отражающая региональную принадлежность ОЭЗ, содержит в обобщенном виде факторы, перечисленные в работе [1], что косвенно может свидетельствовать о влиянии исторических и культурных особенностей на эффективность функционирования зон в РФ. Наше исследование необходимо, чтобы разграничить данные эффекты.

¹ Федеральный закон «Об особых экономических зонах в Российской Федерации» от 22.07.2005 № 116-ФЗ. URL: http://www.consultant.ru/document/cons_doc_LAW_54599

При вступлении фирмы в особую экономическую зону управляющая компания может предоставить отчет о показателях эффективности деятельности ее резидентов, что на самом деле не является гарантом достижения потенциальным резидентом зоны тех или иных экономических и финансовых результатов. Вследствие чего, необходима разработка модели, которая позволила бы выделить индивидуальные эффекты, оказывающие статистически значимое влияние на итоговые финансовые результаты фирмы, рассматривающей возможность

присоединения или неприсоединения к территории зоны. С другой стороны, интерес в разработке модели существует и у управляющей компании, желающей выявить условия, способствующие эффективному процессу генерации резидентов и процессу привлечения инвестиций.

Целью исследования является разработка модели индивидуального бинарного выбора фирмы, перед которой стоит вопрос о возможности вступления или невступления на территорию особой экономической зоны, а также оценка факторов, оказывающих влияние на процесс генерации резидентов и формирование финансового результата действующих резидентов.

Таким образом, предлагаемый в работе подход позволяет дать ответы на вопросы, которые интересуют как резидентов, так и управляющих компаний особых экономических зон.

Гипотеза исследования H1 заключается в том, что фирмы, ведущие деятельность на территории особой экономической зоны, превосходят независимых участников рынка, которые расположены вне территории такой зоны, по величине генерируемой выручке.

Гипотеза исследования H2 относится к выявлению зависимости между решением о вступлении в ОЭЗ и ее местоположением: более привлекательной для потенциальных резидентов может оказаться ОЭЗ, расположенная вблизи центров экономической активности (городов, столиц субъектов с высокой численностью населения), нежели ОЭЗ, размещенные в сельской местности или пригородных территориях.

Структура статьи. Раздел 2 посвящен обзору существующих исследований по темам анализа деятельности ОЭЗ, сетевого взаимодействия, разработки моделей бинарного выбора, алгоритмов машинного обучения. В разделе 3 приводится описание используемых источников данных, способа их сбора

и применяемых методов анализа зависимостей. Спецификация моделей и оценка коэффициентов произведены в разделе 4. В разделе 5 приводится обсуждение факторов, ограничивающих применимость описанных моделей, и озвучиваются потенциальные направления будущих научных изысканий по данной теме. Раздел 6 содержит заключительные положения.

2. Обзор литературы

2.1. Исследования достоинств и недостатков ОЭЗ

С момента учреждения в России в 2005 г. особых экономических зон в качестве одного из наиболее масштабных проектов по привлечению инвестиций в приоритетные секторы экономики большое количество исследований было посвящено анализу эффективности функционирования ОЭЗ, выявлению проблем и предложению перспективных направлений развития.

Радыгина и Суворова [2] приводят обзор особых экономических зон в России, постулируя, что ОЭЗ – эффективный инструмент стимулирования экономического развития и промышленного потенциала страны. Авторы полагают, что налоговые преференции вызовут увеличение темпов генерации резидентов и создадут привлекательные условия для деятельности организаций, послужив «мощным толчком для экономического подъема государства». Проверка гипотезы о наличии такой взаимосвязи и квантификации данного эффекта в исследовании не приводятся.

Иванов и Кожевникова [3] в сравнительном анализе проводят черту между моделями управления особыми экономическими зонами, фокусируясь на региональных отличиях, присущих Дальнему Востоку России. Несмотря на различные подходы к организации экономических пространств в 1990-е гг. и 2010-е гг., авторы указывают на отсутствие фактов, подтверждающих эффективность

улучшения инвестиционной среды в регионе.

Квашнина [4] также критически подходит к вопросу об оценке влияния деятельности резидентов ОЭЗ и заключает, что ухудшение геополитической обстановки оказывает негативное влияние на процесс привлечения как иностранных, так и российских инвестиций, что противоречит целям, ради которых ОЭЗ организуются.

Ахмедзянов и др. [5] обосновывают, что предоставляемые резидентам налоговые льготы в России не являются значительными в сравнении с мировой практикой. Однако эти льготы выступают главным конкурентным преимуществом и помогают снизить первоначальные издержки инвесторов.

Какаулина & Горлов [6] использовали однофакторные степенные регрессионные модели влияния логарифма объема используемых резидентами ОЭЗ РФ льгот на эндогенные переменные: объем текущих инвестиций и CFROI. Авторы доказали наличие положительного воздействия налоговых льгот на инвестиционную деятельность резидентов, что свидетельствует о значимости организации в РФ территорий с преференциальными режимами. При этом эффективность налоговой политики на территориях с особыми режимами ведения предпринимательской деятельности может варьироваться, что связано с различиями в объектах исследования (при выборе конкретных ОЭЗ), периода исследования и применяемых методов.

Min & Kang [7] обсуждают возможность преодоления социо-экономических проблем, которые присущи субъектам Дальнего Востока, путем создания территорий опережающего развития. Авторы утверждают, что в краткосрочном периоде эффект от такого предложения выражен слабо, однако в среднесрочной и долгосрочной перспективе имеются все предпосылки для формирования новых

цепочек создания стоимости, в случае если такие территории докажут свою результативность.

Гуляева и др. [8] обосновывают, что по состоянию на начало 2023 г. проект ОЭЗ стал выходить на траекторию самокупаемости. Сумма перечисленных в бюджет налогов, страховых платежей и таможенных пошлин составила около 303 млрд руб., тогда как сумма полученных резидентами льгот составила около 102 млрд руб.

Бухарова и др. [9] анализируют динамику и структуру ОЭЗ в России. Их число увеличилось с 42 до 45 в 2021. В разрезе типов зон прирост произошел в ОЭЗ промышленно-производственного типа – были созданы 3 новые зоны; технико-внедренческие, туристско-рекреационные и портовые зоны не показали прироста за год.

Корева и др. [10] сравнивают показатели эффективности различных типов ОЭЗ, функционирующих на территории России. Они выделяют зоны промышленно-производственного и технико-внедренческого типов в качестве наиболее успешных по числу генерируемых резидентов. Авторы указывают на важную компоненту – временной период существования зоны, а также аргументируют, что излишняя регламентация и разделение ОЭЗ по типам могут служить сдерживающим фактором, выступая барьером в реальных производственных, технологических и исследовательских цепочках.

Кузнецова [11] указывает на риски невысокой эффективности вложений в НИОКР. Как заключает автор, это не означает, что менее развитые регионы должны быть оставлены в тени – необходимы механизмы сокращения дифференциации.

Доржиева и др. [12] аргументируют наличие дифференцированных подходов для повышения уровня инновационного потенциала и обеспечения социально-экономического развития

в условиях высокой дифференциации регионов страны.

Следует заметить, что подходам, используемым в российских научных исследованиях по вопросам анализа эффективности деятельности ОЭЗ, не хватает некоторой конкретики относительно детерминации силы и направления влияния различных факторов, оказывающих как положительные, так и отрицательные эффекты на результативность таких особых зон в вопросах ведения предпринимательской деятельности.

2.2. Исследования сетевого взаимодействия в ОЭЗ

Chaudey et al. [13] выявили взаимосвязи между местоположением организации и получением ею конкурентных преимуществ, которые ослабевают по мере удаления компании от центра экономической активности города или агломерации. Авторы отмечают, что издержки вступления и поддержания сети могут оказать губительный эффект на индивидуальном уровне для показателей выручки, не принося компании дополнительных доходов в сравнении с независимыми участниками рынка. Стоит отметить, что такой результат может быть обусловлен смещением, возникающим из-за большего количества наблюдений для фирм, расположенных на удаленных территориях.

Неретина [14] указывает на особенности ведения предпринимательской деятельности резидентами ОЭЗ, выделяя те или иные компоненты сетевого взаимодействия, которые, как она предполагает, должны оказать положительный эффект на темпы экономического роста ОЭЗ.

Если предположить, что вступление фирмы в особую экономическую зону можно рассматривать как форму сетевого взаимодействия, где противопоставляются альтернативы: действовать на рынке вне зоны или примкнуть к экономической зоне, которая позволяет

резидентам получить совместный доступ к клубным благам (таким как льготные таможенные и налоговые режимы, объекты инфраструктуры особой экономической зоны, включая генерирующие объекты, транспортные и логистические узлы), то такое сетевое взаимодействие позволяет также получить выгоду новым резидентам благодаря положительному эффекту от привлечения потребителей существующими резидентами ОЭЗ.

Модель, описывающая выгоды и издержки от такого сетевого взаимодействия, в общем виде рассмотренная у Blois [15], может быть описана в терминах трансакционных и трансформационных издержек, возникающих при работе фирмы на и вне территории свободной экономической зоны:

$$VA_{SEZ} - VA_m + TC_m > TC_{SEZ} + (PC_{SEZ} - PC_m), \quad (1)$$

где VA_{SEZ} , VA_m – добавленная стоимость при функционировании фирмы на и вне территории особой экономической зоны соответственно, TC_{SEZ} , TC_m – трансакционные издержки при работе фирмы в составе ОЭЗ и вне нее соответственно, $(PC_{SEZ} - PC_m)$ – разница между трансформационными издержками в ситуации осуществления деятельности фирмы в составе зоны и вне нее.

Можно предположить, что фирма присоединится к особой экономической зоне, если добавленная стоимость, формируемая при осуществлении деятельности в качестве резидента зоны, и экономия на трансакционных издержках, создаваемая благодаря наличию льготных условий на территории зоны, не окажутся меньше величины дополнительных трансформационных и трансакционных издержек, которые фирме необходимо понести, будучи являясь резидентом ОЭЗ.

Проблемы, связанные с таким описанием условий индивидуального принятия решений фирмами о возможности

вступления или невступления в особую экономическую зону, возникают при определении потенциальных величин транзакционных и трансформационных издержек, которые необходимо обосновать. К тому же модель предполагает учет как явных, так и неявных издержек.

В данной работе предлагается подход к разработке модели индивидуального выбора фирмы, основанный на анализе величин, которые явным образом могут быть оценены и которые способны прямо или косвенно учесть предполагаемую величину трансформационных и транзакционных издержек, про которые говорится в обобщенной модели.

2.3. Исследования специфики моделирования потенциальных резидентов

Исследовательский интерес заключается, помимо прочего, в применении современных подходов к решению задачи генерации потенциальных резидентов. Существует необходимость выявления факторов, которые могут сыграть ключевую роль при выборе фирмой одной из альтернатив: войти или нет в состав ОЭЗ в качестве резидента.

Сложность такого моделирования заключается в наличии множества как наблюдаемых явно, так и ненаблюдаемых факторов. Особенностью работы с массивом данных в условиях ограниченного набора информации является проблема пропущенных переменных.

Saar-Tsechansky & Provost [16] обосновали, что многие алгоритмы машинного обучения исходят из предпосылки о 100 % полноте исследуемой выборки, чего в общем виде сложно достичь, имея дело с реальными экономическими данными.

Suthaharan [17] использует метод машинного обучения *Support vector machines* для классификации и детектирования выбросов. Алгоритм позволяет осуществлять бинарную и многоклассовую классификацию. Однако при

значительном превышении числа предикторов над числом доступных наблюдений возникает проблема переобучения.

Rokach & Maimon [18] показывают, что аналогичная проблема свойственна методу *Decision Trees*, который относится к тому же классу методов машинного обучения – обучение с учителем (от англ. supervised learning). Метод использует дерево решений, применяя простые правила «если-то-иначе», с целью предсказания значений целевой переменной. К преимуществам данного метода относят нетребовательность к данным – необходимо произвести нормирование, создание фиктивных переменных и исключение пропущенных значений. Метод использует множество кусочно-постоянных аппроксимирующих функций, которые сложно применять в целях экстраполяции.

Murphy [19] показывает возможность методов класса *Naive Bayes*, основанных на теореме Байеса с предположением об условной независимости между парами признаков. К преимуществам алгоритма можно отнести его более высокую скорость работы, необходимость небольшой обучающей выборки, на основании которой происходит расчет необходимых для модели параметров. Однако вычисление параметров модели предполагает наличие полной информации о значениях предикторов из всего поля признаков.

Mucherino et al. [20] аргументирует перспективы использования алгоритма *K-Nearest Neighbors*, идея которого заключается в нахождении предопределенного числа наблюдений, ближайших по расстоянию к предсказываемому, и определении класса наблюдения на их основе. В случае если в выборке существуют значения с пропущенными переменными, расстояние между наблюдениями искажается, что приводит к неверной классификации.

Peter et al. [21] используют модели *Gradient Boosting*, которые относятся

к классу комплексных моделей в машинном обучении. Они представляют собой комбинацию моделей, которые используются для построения оптимальной предсказательной модели. Такие модели обладают высокой предсказательной точностью и эффективностью. Данные модели применяются для решения регрессионных и классификационных задач. *Gradient Boosting* не поддерживает выборки, в которых отсутствуют значения переменных.

Chaudhary et al. [22] анализируют модель *Random Forest*, которая также относится к комплексным моделям. Если *Decision Trees* генерирует цепочку предсказаний с использованием правил «если-то-иначе», то *Random Forest* объединяет множество таких цепочек с включением случайности, тогда предсказание такого комплекса индивидуальных классификаторов – среднее из предсказанных значений.

Ke et al. [23] предлагают использовать алгоритм классификации *Histogram-based Gradient Boosting Classification Tree*. Он демонстрирует несколько главных преимуществ, среди них преимущества, унаследованные от *Gradient Boosting*, устранение необходимости исключения пропущенных переменных и увеличение скорости классификации алгоритмом.

Tamim Kashifi & Ahmad [24] обосновывают, что, несмотря на всеобщую популярность такого алгоритма классификации, как градиентный бустинг, в машинном обучении, при его применении могут возникнуть проблемы, связанные с пропущенными переменными и выборками большей размерности. Marlin [25] придерживается идентичного мнения относительно возможных проблем с пропущенными переменными.

В данной работе предлагается использовать подход с применением *Histogram-based Gradient Boosting Classification Tree* (HistGBCT). HistGBCT, основанный на градиентном бустинге, задействует меньше вычислительных

мощностей и позволяет не прибегать к методу импутации данных, как показывает исследование Guryanov [26].

3. Данные и методы

3.1. Автоматизированный сбор данных о резидентах ОЭЗ России

При проведении исследования автор столкнулся с проблемой отсутствия выборки необходимых размеров, содержащих нужные для анализа характеристики организаций и ОЭЗ в целом. Для устранения такого пробела и получения достаточной для проведения исследования информации были применены инструменты сбора данных из открытых источников.

Информация об особых экономических зонах на территории России и действующих резидентах была получена на официальном портале Министерства экономического развития РФ¹. Информация в разрезе резидентов зон о состоянии активов и пассивов организаций, данные о финансовых результатах деятельности фирм были получены с помощью Государственного информационного ресурса бухгалтерской (финансовой) отчетности (ресурс БФО)².

Сведения о состоянии организаций, основном виде деятельности, юридическом адресе, среднесписочной численности работников получены на портале сервиса «Прозрачный бизнес»³. Для дополнения выборки сведениями о динамике развития ОЭЗ использовались данные из Бизнес-навигатора по особым экономическим зонам России – 2022 [8].

¹ Особые экономические зоны. Министерство экономического развития Российской Федерации. URL: https://www.economy.gov.ru/material/directions/regionalnoe_razvitiye/instrumenty_razvitiya_territoriy/osoby_e_ekonomicheskie_zony

² Государственный информационный ресурс бухгалтерской (финансовой) отчетности. URL: <https://bo.nalog.ru>

³ Сервис «Прозрачный бизнес». URL: <https://pb.nalog.ru/index.html>

Итоговая выборка данных по резидентам ОЭЗ РФ включает 124 кросс-секционных наблюдения за 17 временных периодов с 2005 по 2021 г. В работе было рассмотрено 2 типа ОЭЗ – портовые и промышленно-производственные.

3.2. Балансирование выборки: простой случай отбор предприятий-нерезидентов

Кроме того, для анализа различий между предприятиями, не являющимися резидентами ОЭЗ, выборка была дополнена информацией о 353 предприятиях за аналогичный временной период, которая была получена в ходе простого случайного отбора ИНН российских организаций, действовавших на момент проведения исследования.

3.3. Итоговая выборка резидентов и нерезидентов

Совокупная выборка включает такие показатели, как ИНН; название фирмы; ОКВЭД; статьи бухгалтерского баланса и отчета о финансовых результатах; а также рассчитанные на их основе показатели рентабельности (коэффициент рентабельности продаж (ROS), коэффициент рентабельности собственного капитала (ROE), коэффициент рентабельности активов (ROA)); коэффициент автономии; коэффициенты текущей и промежуточной ликвидности; бинарную переменную, которая показывает, является ли фирма резидентом ОЭЗ; год учреждения организации; юридический адрес; среднесписочную численность работников.

Подвыборка резидентов ОЭЗ также включает такие переменные, как количество резидентов в динамике, динамика частных инвестиций, размер уплаченных налогов и налоговых льгот, количество созданных рабочих мест, производительность территории, а также информацию об объектах инфраструктуры ОЭЗ (площадь, мощность объектов и стоимость водо-, электро-, газо-, теплоснабжения, протяженность автодорог).

3.4. Регрессионные и классификационные инструменты

Для оценки предполагаемых зависимостей и тестирования выдвинутых гипотез применялись инструменты корреляционно-регрессионного анализа и методы классификации в машинном обучении.

К первой группе использованных инструментов относится регрессионная модель панельных данных с фиксированными и случайными эффектами. Ко второй группе относится первый ряд инструментов, требующих ресемплинга данных и исключения пропущенных переменных, алгоритмы классификации *Support Vector Machines*, *Decision Trees*, *Random Forest*, *Naive Bayes*, *K-Nearest Neighbor*, *Gradient Boosting*, *Logistic Regression*, а также второй ряд – алгоритм классификации, устойчивый для работы с большими выборками и имеющий встроенную поддержку данных с пропущенными переменными, HistGBCT.

3.5. Подготовка данных

При проведении регрессионного анализа панельных данных были исключены линейно-зависимые факторы, наличие которых может привести к мультиколлинеарности в модели. На основании VIF-теста, который рекомендуют использовать Mansfield & Helms [27] при спецификации моделей панельных данных со случайными и фиксированными эффектами, исключались факторы, для которых значения превышали ⁴1.

Проведенный тест на гомоскедастичность остатков модели показал, что нулевая гипотеза о гомоскедастичности отвергается в пользу альтернативной,

¹ Variance Inflation Factor (VIF). Corporate Finance Institute. URL: <https://corporatefinanceinstitute.com/resources/data-science/variance-inflation-factor-vif>

таким образом, в дальнейшем в работе используются робастные стандартные ошибки. По мнению Kleiber & Zeileis [28], это позволит повысить надежность получаемых оценок коэффициентов и уменьшить величину стандартных ошибок.

Как было отмечено выше, для работы большинства алгоритмов машинного обучения должна выполняться строгая предпосылка об отсутствии пропущенных переменных. В связи с этим пропущенные переменные во время работы алгоритмов из второй группы первого ряда были заполнены средними значениями по выборке.

Также проводилась нормировка данных для повышения скорости работы алгоритмов, увеличения вероятности схождения функции оптимизации и снижения чувствительности алгоритмов к единицам измерения данных. Применяемый метод нормировки

(*MinMaxScaler*) описан в работе Деера & Рамеш [29].

4. Результаты

4.1. Описательные статистики выборки

На рис. 1 представлена карта Российской Федерации, сформированная с помощью библиотеки *folium*¹ 0.14.0, на которой отмечены организации, отобранные в ходе простого случайного отбора, составляющие подвыборку независимых фирм, т. е. не являющихся резидентами ОЭЗ.

В табл. 2 приведены описательные статистики по всей исследуемой выборке, а также в разрезе компаний-резидентов ОЭЗ и нерезидентов. Естественным образом наблюдаются значительные различия полученных значений для выбранных переменных из поля факторов для организаций, входящих и не входящих в ОЭЗ.

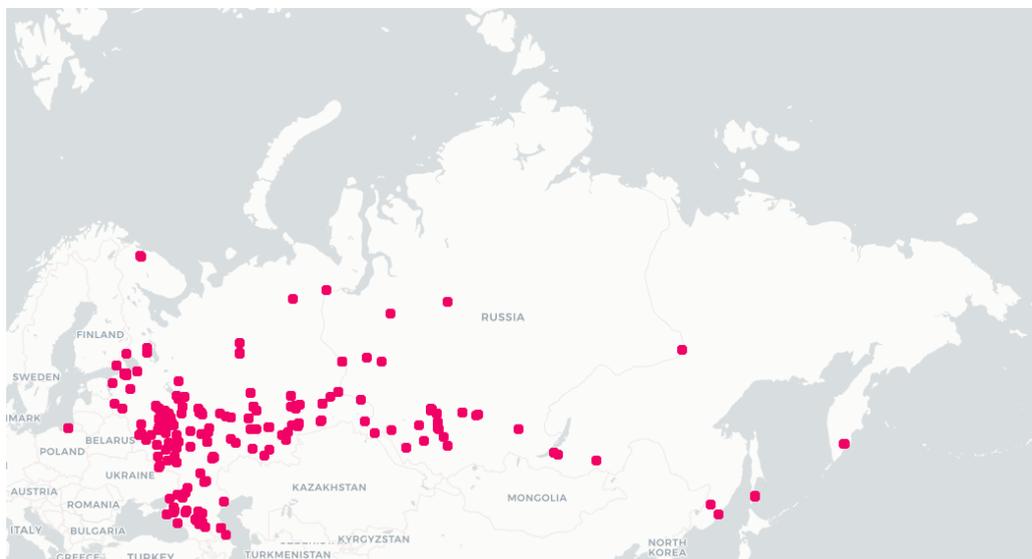


Рис. 1. Карта РФ с изображением независимых организаций, отобранных в ходе простого случайного отбора ИНН российских компаний

Fig. 1. Map of the Russian Federation depicting independent organizations selected during a simple random sampling of the Russian companies' TINs

Источник: составлено автором на основании данных из ЕГРЮЛ/ЕГРИП²

¹ Folium. The Python Package Index. URL: <https://pypi.org/project/folium>

² Предоставление сведений из ЕГРЮЛ/ЕГРИП. URL: <https://egrul.nalog.ru/index.html>

Таблица 2. Сводная матрица описательных статистик

Table 2. Summary matrix of descriptive statistics

Переменная	№ наблюдений	Среднее	Ст. откл.	Медиана	Минимум	Максимум
<i>По всей выборке</i>						
Категориальная переменная местоположения	7632	2,30	0,88	3,0	1,00	3,00
Расстояние до Москвы, км	7632	901,21	1094,27	633,01	0,56	6797,20
Расстояние до центра субъекта, км	7632	62,01	98,51	20,00	0,00	1121,43
Выручка, млрд руб.	7632	10,82	52,10	0,33	0,00	792,93
Чистая прибыль, млрд руб.	7632	1,00	10,17	0,00	-96,53	278,19
Коэффициент оборачиваемости активов	7632	288,19	9793,99	0,78	0,00	404099,27
Коэффициент автономии	7632	-3,80	147,56	0,47	-6247,85	32,49
Коэффициент текущей ликвидности	7632	29,08	349,73	1,51	0,00	12024,71
ROS	7632	-1,77	95,13	0,20	-3911,00	1022,88
ROA	7632	-0,43	15,75	0,01	-660,80	12,84
ROE	7632	0,47	13,02	0,07	-47,55	760,97
<i>Резиденты ОЭЗ</i>						
Категориальная переменная местоположения	1984	1,73	0,83	1,00	1,00	3,00
Расстояние до Москвы, км	1984	777,80	446,65	986,00	73,00	2720,00
Расстояние до центра субъекта, км	1984	99,90	79,39	73,00	0,00	213,40
Выручка, млрд руб.	1984	1,21	2,40	0,44	0,00	26,49
Чистая прибыль, млрд руб.	1984	-0,02	0,52	0,00	-2,88	9,42
Коэффициент оборачиваемости активов	1984	1,10	1,70	0,59	0,00	20,75
Коэффициент автономии	1984	0,35	0,95	0,29	-9,98	17,72

Окончание табл. 2

Переменная	№ наблюдений	Среднее	Ст. откл.	Медиана	Минимум	Максимум
Коэффициент текущей ликвидности	1984	47,38	637,36	1,39	0,01	12024,71
ROS	1984	-16,53	227,63	0,17	-3911,00	103,00
ROA	1984	0,03	0,63	0,01	-6,89	12,84
ROE	1984	1,52	28,20	0,11	-25,16	760,97
<i>Организации, не входящие в ОЭЗ</i>						
Категориальная переменная местоположения	5648	2,46	0,83	3,00	1,00	3,00
Расстояние до Москвы, км	5648	934,43	1208,71	601,55	0,56	6797,20
Расстояние до центра субъекта, км	5648	51,81	100,65	10,80	0,00	1121,43
Выручка, млрд руб.	5648	12,72	56,82	0,30	0,00	792,93
Чистая прибыль, млрд руб.	5648	1,23	11,23	0,00	-96,53	278,19
Коэффициент оборачиваемости активов	5648	344,95	10717,86	0,81	0,00	404099,27
Коэффициент автономии	5648	-4,78	164,09	0,50	-6247,85	32,49
Коэффициент текущей ликвидности	5648	24,92	240,28	1,54	0,00	5388,38
ROS	5648	1,10	26,28	0,20	-48,02	1022,88
ROA	5648	-0,54	17,41	0,02	-660,80	4,23
ROE	5648	0,23	5,17	0,06	-47,55	210,00

Источник: рассчитано автором.

Кроме того, проведенный тест по непараметрическому критерию суммы рангов Уилкоксона, апробированный Fay & Proschan [30], согласно методологии, представленной Ng et al. [31], показал, что две независимые выборки на примере уровня переменных ROS, ROE, ROA действительно взяты из различных распределений, нулевая гипотеза отвергается в пользу альтернативной. Достаточно, что хотя бы для 1 признака нулевая гипотеза была отвергнута

для того, чтобы заключить о различиях между анализируемыми подвыборками.

В табл. 3 приведено частотное распределение видов экономической деятельности для организаций, входящих в состав ОЭЗ, и для организаций, действующих за пределами ОЭЗ.

Исходя из анализа видов экономической деятельности, можно сделать вывод о том, что резиденты, с учетом законодательных ограничений, накладываемых на допустимые виды деятельности

Таблица 3. **Частотное распределение видов экономической деятельности**
 Table 3. **Frequency distribution of economic activities**

№	ОКВЭД	Выборочная частота
<i>Организации, не входящие в ОЭЗ</i>		
1	Аренда и управление собственным или арендованным нежилым недвижимым имуществом (68.20.2)	243
2	Разведение молочного крупного рогатого скота, производство сырого молока (1.41)	94
3	Передача электроэнергии и технологическое присоединение к распределительным электросетям (35.12)	65
4	Торговля оптовая неспециализированная (46.90)	64
5	Вложения в ценные бумаги (64.99.1)	55
<...>		
219	Деятельность в области архитектуры (71.11)	5
220	Торговля оптовая станками (46.62)	4
221	Производство стальных труб, полых профилей и фитингов (24.20)	4
222	Ремонт часов и ювелирных изделий (95.25)	3
223	Производство вертолетов, самолетов и прочих летательных аппаратов (30.30.3)	3
<i>Резиденты ОЭЗ</i>		
1	Производство прочих комплектующих и принадлежностей для автотранспортных средств (29.32)	53
2	Деятельность по складированию и хранению (52.10)	43
3	Производство бумажных изделий хозяйственно-бытового и санитарно-гигиенического назначения (17.22)	33
4	Производство прочих химических продуктов, не включенных в другие группировки (20.59.5)	31
5	Производство пластмассовых плит, полос, труб и профилей (22.21)	29
<...>		
86	Ремонт и техническое обслуживание летательных аппаратов, включая космические (33.16)	3
87	Производство пиломатериалов, профилированных по кромке; производство древесного полотна, древесной муки; производство технологической щепы или стружки (16.10.2)	3
88	Производство прочей одежды и аксессуаров одежды (14.19)	3
89	Производство спецодежды (14.12)	3
90	Производство прочих химических органических основных веществ (20.14.7)	2

Источник: составлено автором.

в пределах ОЭЗ, специализируются чаще всего на выпуске продукции с высокой добавленной стоимостью. Наличие такой специализации аргументируют Дебердиева и др. [32].

Табл. 4 содержит описательные статистики в разрезе типов ОЭЗ, исследуемых в работе.

Исходя из анализа исследований, посвященных перспективам развития

Таблица 4. **Описательные статистики портовых и промышленно-производственных типов ОЭЗ в России**

Table 4. **Descriptive statistics of port and industrial-production types of SEZs in Russia**

Переменная	Статистика	Портовые ОЭЗ	Промышленно-производственные ОЭЗ	Переменная	Портовые ОЭЗ	Промышленно-производственные ОЭЗ
Категориальная переменная местоположения	Среднее	2,00	1,71	Объем налоговых льгот в 2020 г., млн руб.	0,54	0,06
	Ст. откл.	0,00	0,85		0,00	0,13
	Минимум	2,00	1,00		0,54	0,00
	Медиана	2,00	1,00		0,54	0,00
	Максимум	2,00	3,00		0,54	0,75
Расстояние до Москвы км	Среднее	875,00	771,59	Выручка, млрд руб.	0,22	1,27
	Ст. откл.	0,00	460,02		0,51	2,45
	Минимум	875,00	73,00		0,00	0,00
	Медиана	875,00	986,00		0,06	0,50
	Максимум	875,00	2720,00		3,01	26,49
Расстояние до центра субъекта, км	Среднее	35,00	104,04	Чистая прибыль, млрд руб.	0,03	-0,02
	Ст. откл.	0,00	80,12		0,10	0,53
	Минимум	35,00	0,00		-0,02	-2,88
	Медиана	35,00	100,00		0,00	0,00
	Максимум	35,00	213,40		0,58	9,42
Количество созданных рабочих мест на 2020 г., ед.	Среднее	696,00	3863,81	Коэффициент оборачиваемости активов	0,93	1,11
	Ст. откл.	0,00	2740,89		0,89	1,73
	Минимум	696,00	0,00		0,00	0,00
	Медиана	696,00	4694,00		0,74	0,59
	Максимум	696,00	7287,00		2,90	20,75
Средняя производительность труда территории в 2020 г., млн руб./чел.	Среднее	2,20	8,23	Коэффициент автономии	0,27	0,36
	Ст. откл.	0,00	3,56		0,66	0,96
	Минимум	2,20	0,00		-1,98	-9,98
	Медиана	2,20	6,70		0,23	0,29
	Максимум	2,20	12,00		1,21	17,72

Окончание табл. 4

Переменная	Статистика	Портовые ОЭЗ	Промышленно-производственные ОЭЗ	Переменная	Портовые ОЭЗ	Промышленно-производственные ОЭЗ
Протяженность автомобильных дорог, м	Среднее	6994,00	22199,20	Коэффициент текущей ликвидности	3,83	49,82
	Ст. откл.	0,00	13115,06		6,74	654,90
	Минимум	6994,00	1093,00		0,13	0,01
	Медиана	6994,00	27757,00		1,70	1,38
	Максимум	6994,00	37887,00		38,52	12024,71
Среднегодовой темп прироста резидентов, %	Среднее	27,54	20,90	ROS	-3,04	-17,33
	Ст. откл.	0,00	15,86		11,88	234,28
	Минимум	27,54	8,84		-67,50	-3911,00
	Медиана	27,54	14,94		0,24	0,17
	Максимум	27,54	74,11		1,00	103,00
Среднегодовой темп прироста инвестиций, %	Среднее	109,65	16,49	ROA	0,03	0,03
	Ст. откл.	0,00	45,90		0,34	0,65
	Минимум	109,65	-14,68		-1,60	-6,89
	Медиана	109,65	-13,72		0,03	0,01
	Максимум	109,65	213,42		0,56	12,84
Объем уплаченных налогов в 2020 г. в расчете на резидента, млн руб.	Среднее	3,17	29,91	ROE	3,55	1,40
	Ст. откл.	0,00	30,02		18,40	28,64
	Минимум	3,17	-32,11		-6,52	-25,16
	Медиана	3,17	24,36		0,36	0,10
	Максимум	3,17	70,71		112,80	760,97

Источник: рассчитано автором исследования.

и проблемам экспансии ОЭЗ в России, проведено тестирование гипотезы о наличии взаимосвязи между уровнем социально-экономического развития регионов и процессом генерации резидентов и частных инвестиций. Соответствующая корреляционная матрица, включающая данные по трем типам ОЭЗ (промышленно-производственным, технико-внедренческим, портовым), приведена на рис. 2.

Для обнаружения возможной связи использованы такие показатели, как

валовой региональный продукт на душу населения¹, инвестиции в основной капитал по полному кругу организаций² и среднедушевые денежные доходы

¹ Валовой региональный продукт на душу населения. Витрина статистических данных. URL: <https://showdata.gks.ru/olap2/descr/report/277524> (дата обращения: 05.01.2023).

² Инвестиции в основной капитал по формам собственности по полному кругу организаций. Витрина статистических данных. URL: <https://showdata.gks.ru/olap2/descr/report/274074> (дата обращения: 05.01.2023).

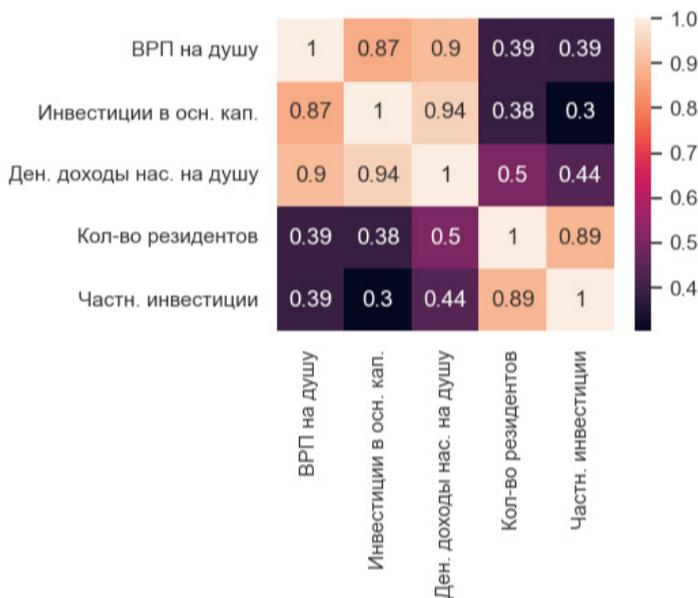


Рис. 2. Корреляционная матрица, описывающая силу связи между переменными пространственной дифференциации регионов и процессом генерации резидентов и частных инвестиций в ОЭЗ России

Fig. 2. Pairwise correlation matrix between variables of regions spatial differentiation vs. resident and private investment generation process in SEZs

Источник: составлено автором на основании [8].

населения¹ за 2020 г., согласно методике, рассматриваемой в работах Benini & Czyzewski [33], Yershov [34].

Рис. 3. демонстрирует графики разброса для соответствующих показателей. В целом связи могут быть интерпретированы как умеренно положительные (Schober et al. [35]).

4.2. Корреляционно-регрессионный анализ процесса генерации резидентов

В данном разделе представлена регрессионная модель панельных данных, описывающая влияние факторов на средний темп прироста резидентов ОЭЗ. Исследуемой переменной выступает среднегодовой темп прироста резидентов ОЭЗ. При построении данной

модели используется только подвыборка предприятий, входящих в состав ОЭЗ России, поскольку по определению для независимых предприятий среднегодовой темп прироста резидентов не может быть указан.

На рис. 4 приведены динамические ряды, отражающие среднегодовые темпы прироста резидентов и частных инвестиций за период 2013–2021 гг. в разрезе двух типов ОЭЗ – портовых (ПОЭЗ) и промышленно-производственных (ОЭЗ ППТ).

Для построения множественной регрессионной модели панельных данных используется метод отбора факторов, основанный на анализе факторов инфляции дисперсии. Критическим значением в данной работе принимается величина, равная 4, из всего поля факторов отбирается максимальное число регрессоров, позволяющих получить линейно-независимую комбинацию. Возможные спецификации множественной регрессионной

¹ Денежные доходы (в среднем на душу). Единая межведомственная информационно-статистическая система. URL: <https://www.fedstat.ru/indicator/30992> (дата обращения: 05.01.2023).

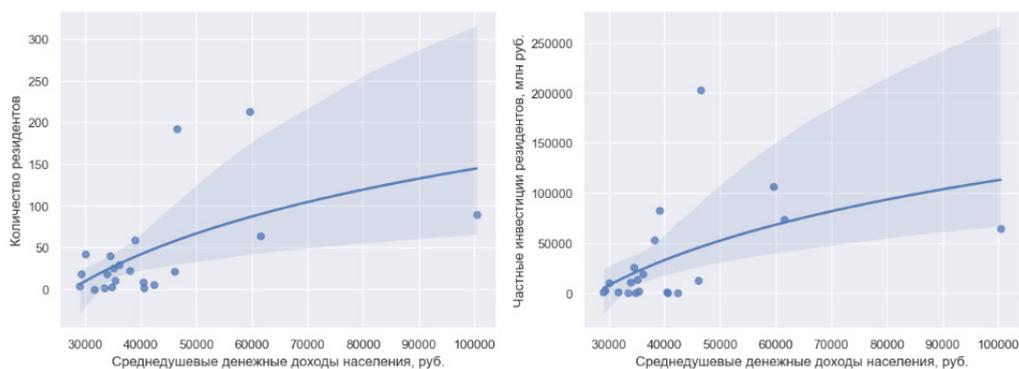


Рис. 3. Графики разброса, описывающие взаимосвязь между процессом генерации резидентов и частных инвестиций в ОЭЗ и величиной среднедушевых денежных доходов населения по субъектам РФ за 2020 г.

Fig. 3. Scatter plots depicting resident and private investment generation process vs. monetary income per capita across the regions of Russia in 2020.

Источник: составлено автором.

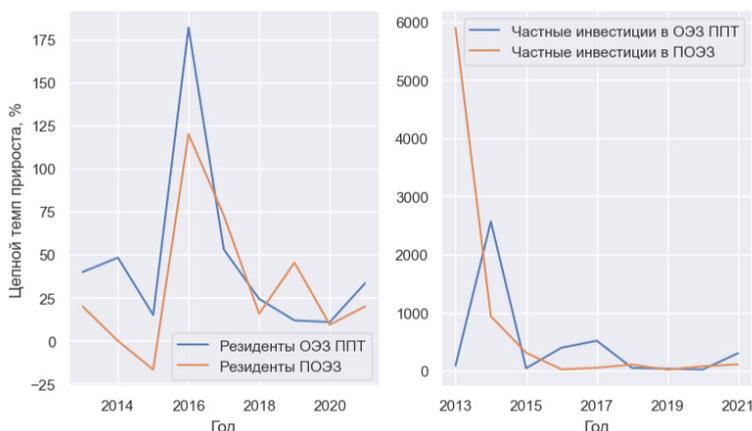


Рис. 4. График среднегодового темпа прироста резидентов и частных инвестиций в ОЭЗ ППТ и ПОЭЗ за период 2013–2021 гг.

Fig. 4. Average annual growth rate of residents and private investments in industrial-production and port types of SEZs, 2013–2021

Источник: составлено автором на основании [8].

модели панельных данных со случайными эффектами представлены в табл. 5.

В работе было рассмотрено три возможных спецификации модели.

Переменная «Протяженность автомобильных дорог» отражает абсолютную длину, которую транспортные сети занимают на территории ОЭЗ. Увеличение протяженности автомобильных дорог ведет к сокращению среднегодового темпа прироста резидентов, что имеет прямые физические следствия

из-за ограниченности территории ОЭЗ, устанавливаемой законодательством. Объем уплаченных налогов в расчете на резидента также оказывает негативное влияние на зависимый признак, поскольку является элементом структуры затрат организации, осуществляющей предпринимательскую деятельность. Переменная «местоположения» является категориальной переменной, принимающей значения из множества {1, 2, 3}, где 1 – отражает удаленность

Таблица 5. Спецификации множественной регрессионной модели, отражающей влияние регрессоров на процесс генерации резидентов в ПОЭЗ и ОЭЗ ППТ

Table 5. Econometrical estimations for the average annual growth rate of residents

Зависимая переменная: среднегодовой темп прироста резидентов, % Кросс-секционные наблюдения: 124 организации Временной период: 2012–2021 гг.			
Регрессор	1	2	3
const	39,508* (1,2978)	23,229* (1.1032)	28,380* (1.0023)
Протяженность автомобильных дорог, м	-0,0006* (5,425e-05)	-0,0004* (1,131e-05)	-0,0005* (8,601e-06)
Объем уплаченных налогов резидентами в расчете на 1 резидента, млн руб.	-0,2479* (0,0231)	-0,1403** (0,0061)	–
Местоположение	1,6133** (0,7253)	–	–
Тип ОЭЗ	–	2,5476* (0,2010)	2,3549* (0,1887)
Объем налоговых льгот в расчете на 1 резидента, млн руб.	–	3,2066** (0,1027)	4,7946* (0,1239)
Расстояние до центра субъекта, км	–	–	-0,0526* (0,0013)
Код ОКВЭД	–	–	0,0100 (0,0082)
Код ОЭЗ	–	2,4405* (0,1900)	–
Код региона	0,1578** (0,0005)	–	–
Мах (VIFi)	3,23	2,36	2,22
Стандартные ошибки	HWse ¹	HWse	HWse
R ²	0,7249	0,9077	0,9281
F-статистика	980,07	1809,3	2257,1

Источник: составлено автором исследования.

Примечание: в скобках под значениями оцененных коэффициентов указаны стандартные ошибки. * – значимость коэффициентов указана на 1 %-м уровне, ** – значимость коэффициентов указана на 5 %-м уровне, *** – значимость коэффициентов указана на 10 %-м уровне.

¹ В работе используются стандартные ошибки в форме Уайта (HWse) на основании теста на гетероскедастичность остатков модели.

фактического местоположения организации от центра субъекта более чем на 80 км, 2 – в пределах от 35 до 80 км и 3 – менее чем на 35 км.

Таким образом, гипотеза о том, что процесс генерации резидентов проходит более успешно вблизи центров экономической активности субъектов и городских агломераций, может быть подтверждена.

«Тип ОЭЗ» – бинарная переменная, принимающая значения 1, если тип экономической зоны промышленно-производственный; 0, если тип портовый. Увеличение объема налоговых льгот в расчете на резидента стимулирует процесс генерации резидентов по абсолютной величине, не считая константы в уравнении, более всего. Переменная, характеризующая расстояние до центра субъекта, по своему смыслу схожа с категориальной переменной «местоположения», с той лишь разницей, что учитывает абсолютный эффект от удаленности фактического местоположения ОЭЗ от столицы субъекта на процесс генерации резидентов, что также служит подтверждением второй гипотезы, сформулированной во введении работы.

«Код ОКВЭД» – бинарная переменная, которая принимает значение 1, в случае если основной вид деятельности фирмы входит в число пяти видов деятельности, наиболее часто встречающихся среди фирм – резидентов ОЭЗ России, 0 – в ином случае. Категориальная переменная «Код ОЭЗ» принимает значения в соответствии со словарем значений: {Ульяновск: 1; Алабуга: 2; Липецк: 3; Тольятти: 4; Титановая долина: 5; Стабна: 6; Моглино: 7; Калуга: 8; Ступино квадрат: 9; Узловая: 10; Лотос: 11; Центр: 12; Орел: 13; Кашира: 14; Кулибин: 15; Алга: 16; Доброград-1: 17; Авангард: 18; Оля: 19} и контролирует эффект от региональных различий, историко-культурных особенностей территорий, на которых расположены ОЭЗ, категориальная переменная «Код региона»

служит для аналогичных целей и принимает значения согласно Классификатору кодов субъектов Российской Федерации.

Спецификация третьей модели имеет наибольшую объясняющую способность с коэффициентом детерминации, равным 92,81%. Все регрессоры, кроме переменной «Код ОКВЭД», значимы на 1%-м уровне. *P*-значение *F*-статистики, стремящееся к нулю, свидетельствует о совместной значимости выбранных регрессоров. Для любого из регрессоров спецификации 3 значение *VIF*-теста не превышает 2,22, что, вероятно, говорит об отсутствии мультиколлинеарности в модели.

4.3. Регрессионная модель панельных данных со случайными эффектами для анализа влияния факторов на финансовые показатели деятельности компаний

В данной модели в качестве зависимой переменной выступает выручка компаний. Производится попытка оценки влияния переменных местоположения, принадлежности компании к ОЭЗ на формируемую величину выручки по итогам финансового года. Результаты оценки представлены в табл. 6.

В данной модели регрессоры «Местоположение», «Тип ОЭЗ», «Код ОКВЭД», «Код региона», «Объем налоговых льгот в расчете на 1 резидента», «Протяженность автомобильных дорог» аналогичны используемым в предыдущем разделе.

Был добавлен ряд факторов, среди них «Network bin» – бинарная переменная, принимающая значение 1 в случае, если компания является резидентом ОЭЗ, 0 – в противном случае. Произведение переменных местоположения и «Network bin», коэффициент регрессора, как ожидается, должен быть положительным, если верна гипотеза о том, что резиденты, расположенные вблизи городских центров экономической активности,

Таблица 6. Спецификации множественной регрессионной модели, отражающей влияние регрессоров на процесс формирования выручки компаний-резидентов и нерезидентов.

Table 6. Econometrical estimations for the residents and independent outlets' revenue

Зависимая переменная: выручка, тыс. руб. Кросс-секционные наблюдения: 477 организации Временной период: 2012–2021 гг.			
Регрессор	1	2	3
const	4,921e+06*** (3,018e+06)	6,552e+06*** (3,467e+06)	4,926e+06** (3,003e+06)
Местоположение	–	1,472e+06** (0,601e+05)	3,341e+06 (1,677e+05)
Тип ОЭЗ	–	–	10823,7 (8967,0)
Код ОКВЭД	6824,1*** (3973,5)	4646,5** (3274,1)	–
Код региона	7,053e+05*** (3,864e+04)	1,293e+05** (7,538e+04)	–
Объем налоговых льгот в расчете на 1 резидента, млн руб.	–	3,97e+05** (5,008e+04)	2,21e+05 (3,121e+05)
Протяженность автомобильных дорог, м	60,994 (38,818)	–	3,5183 (2,5987)
Network bin	2,337e+06** (1,041e+05)	–	3,777e+06*** (2,421e+05)
(Местоположение × Network bin)	–	–	6,244e+06*** (0,834e+05)
Max (VIFi)	3,21	3,12	3,63
Стандартные ошибки	HWse	HWse	HWse
R ²	0,4963	0,4338	0,5311
F-статистика	661,04	587,94	1513,8

Источник: составлено автором исследования.

Примечание: в скобках под значениями оцененных коэффициентов указаны стандартные ошибки. * – значимость коэффициентов указана на 1 %-м уровне, ** – значимость коэффициентов указана на 5 %-м уровне, *** – значимость коэффициентов указана на 10 %-м уровне.

получают максимальный эффект от сетевого взаимодействия благодаря вхождению в состав ОЭЗ.

Модель не включает действительно значимых переменных, оказывающих прямой результат на формирование

выручки компаний, о чем свидетельствует низкое значение коэффициента детерминации.

Идея построения такой регрессии заключалась в выявлении эффектов, связанных с двумя введенными

переменными, однако неустранимое смещение ввиду пропуска значимых переменных, оказывающих влияние на зависимый признак, высокая волатильность оцененных коэффициентов, позволяют говорить о неудовлетворительном результате, требуется разработка новой спецификации модели.

Интерпретации коэффициенты не подвергаются, однако полезным является тот факт, что можно судить о знаке коэффициентов. Как и предполагается, компании-резиденты получают положительный эффект от ведения деятельности на территории ОЭЗ, близость к центрам экономической активности положительно влияет на ведение предпринимательской деятельности, положительный синергетический эффект достигается при совпадении двух данных факторов.

4.4. Логистическая регрессия и классификационные алгоритмы машинного обучения в предсказании решения потенциального резидента о вхождении в состав ОЭЗ

Предсказать решение потенциального резидента о его вхождении в состав ОЭЗ и начале экономической деятельности на ее территории становится возможным благодаря применению калибруемых алгоритмов машинного обучения. Применение большинства алгоритмов требует выполнения ряда предпосылок, что приводит к необходимости трансформации изначальной выборки.

Первая трансформация данных связана с восстановлением пропущенных переменных на основании всего поля факторов путем осреднения ближайших k -наблюдений, данные по которым присутствуют в выборке – множественная импутация. Выполнение предпосылки о наличии равного числа наблюдений по предприятиям, входящим в состав ОЭЗ и не являющихся резидентами зон, требует проведение

ресемплинга. Подвыборка организаций – резидентов ОЭЗ была дополнена сгенерированными на основании имеющихся данных наблюдениями так, что соотношение между двумя подвыборками стало равным 50 и 50 %.

Таким образом, итоговая выборка содержит 11 296 наблюдений, из которых 5 648 относятся к организациям – резидентам ОЭЗ. Данные также были нормированы с использованием MinMaxScaler с целью приведения их к распределению Гаусса с нулевым средним и единичной дисперсией.

Выборка случайным образом была разделена на тренировочную и тестовую часть, размер которой составляет 30 % от изначального размера, с целью проверки точности работы алгоритма после процесса его обучения на тренировочной части датасета. Результаты проверки работы алгоритмов и точности их предсказания отражены на рис. 5.

Метод опорных векторов успешно справляется с задачей бинарной классификации организаций по признаку принадлежности или непринадлежности к ОЭЗ совместно с классификатором «случайный лес», метрики Accuracy, Precision и Recall составили около 0,99.

Построенная логистическая регрессия и метод k -ближайших соседей чуть менее успешно справляются с поставленной задачей выявления организаций-резидентов и нерезидентов. Градиентный бустинг, наивный байесовский классификатор и метод дерева решений в целом по метрике Accuracy, отражающей совокупную точность предсказания моделей по положительному и отрицательному классам, показали результаты с низким процентом верных предсказаний.

Алгоритм классификации HistGBCT в качестве альтернативы градиентному бустингу призван повысить скорость обучения моделей при использовании для больших выборок. Данный классификатор включает нативную поддержку пропущенных переменных, что позволяет

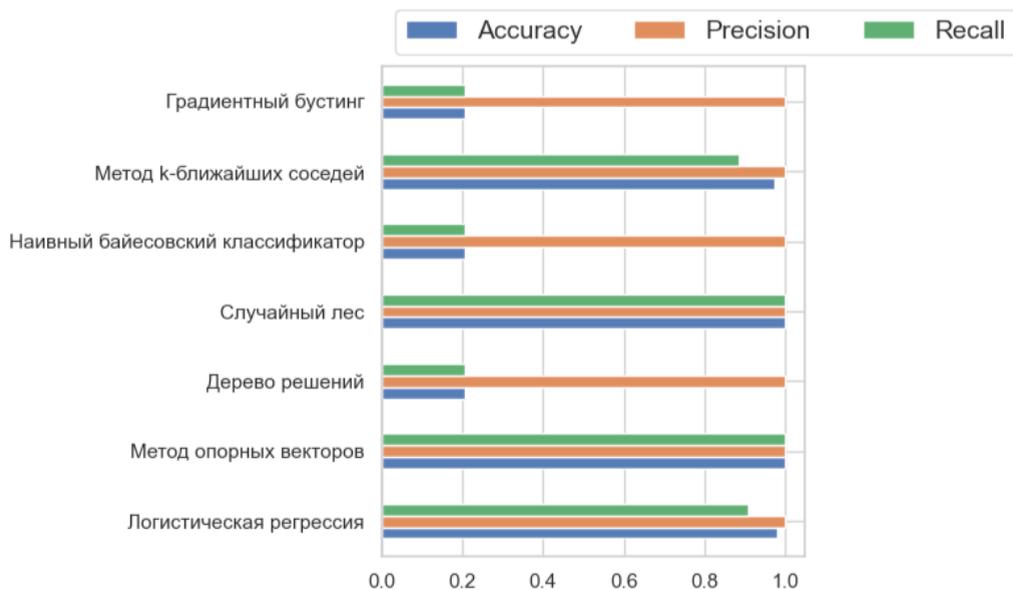


Рис. 5. Сравнение классификационных и регрессионной моделей по точности предсказания

Fig. 5. Comparison of classification and regression models by prediction accuracy

Источник: составлено автором.

исключить необходимость в процедуре предварительного восстановления данных путем, например множественной импутации. За счет значительного прироста скорости работы алгоритма процедура нормировки данных является опциональной и может применяться для повышения вероятности схождения алгоритма.

Выборка была разделена на две части: тренировочную и тестовую. При применении данного алгоритма выборка использовалась в том виде, в котором была получена на стадии сбора информации, импутация, ресемплинг не производились. Оценка работы алгоритма и его прогнозной точности представлены в таблице 7 вместе с значениями коэффициентов и их стандартными ошибками.

Подводя итог раздела по построению модели прогноза вероятности вхождения организации в состав ОЭЗ, необходимо отметить высокую предсказательную точность HistGBCT, кроме того, его широкую применимость, связанную с низкой требовательностью

к данным, над которыми не требуется выполнение ряда трансформаций перед началом работы классификатора.

Данный алгоритм может быть использован для решения ряда классификационных задач при работе с реальными экономическими данными, которые часто могут содержать некоторое количество пропущенных переменных по некоторому подмножеству наблюдений, исключение которых может привести к частичной потере информации.

Стоит также дополнительно отметить, что описанный в работе Ke et al. [23] *Histogram-based Gradient Boosting Regression Tree*, обладает теми же свойствами, что и классификатор, но предназначен для решения регрессионных задач, область применения – большие экономические данные, отличающиеся свойством неполноты.

5. Обсуждение

Полученные оценки коэффициентов в регрессионных моделях панельных данных со случайными эффектами

Таблица 7. Результаты оценки коэффициентов методом HistGBCT

Table 7. HistGBCT coefficients estimation results

Регрессор	Значение коэффициента	Ст. ошибка
Местоположение	0,1486*	0,0120
Расстояние до Москвы, км	-0,0001*	0,000001
Расстояние до центра субъекта, км	-0,0040*	0,0003
Средняя производительность труда территории, млн руб./чел.	0,0021***	0,0009
Код региона	-0,0038*	0,0005
Мощность объектов теплоснабжения, Гкал/ч	0,0017***	0,0005
Стоимость электроснабжения, руб./кВт*ч	-0,0017***	0,0010
Объем уплаченных налогов резидентами в расчете на 1 резидента, млн руб.	-0,0981*	0,0042
Коэффициент отношения полученных льгот к величине уплаченных налогов резидентом	0,0934**	0,0293
Коэффициент автономии	0,0008**	0,0005
ROS	0,0732*	0,0173
ROE	0,0071**	0,0017
Тип ОЭЗ	0,4530*	0,0146
Метрика		Значение
Accuracy		0,98
Precision		0,97
Recall		0,99
Коэффициент детерминации прогноза		0,95

Источник: рассчитано автором.

Примечание: * – значимость коэффициентов указана на 1 %-м уровне, ** – значимость коэффициентов указана на 5 %-м уровне, *** – значимость коэффициентов указана на 10 %-м уровне.

позволяют подтвердить сформулированные в начале исследования гипотезы.

Так, переменная, контролирующая принадлежность организации к той или иной ОЭЗ, принимающая в таком случае значение, равное единице, оказалась значимой в моделях на соответствующих уровнях значимости.

Кроме того, положительно оцененный коэффициент свидетельствует о наличии положительного эффекта на показатели финансового результата

деятельности фирм-резидентов; в случае если компания не является резидентом, переменная принимает значение 0, положительный эффект нивелируется.

Так, было доказано, что ОЭЗ действительно оказывают положительные эффекты, связанные с предоставлением резидентам клубных благ, в частности налоговых преференций, доступа к инфраструктурным объектам и транспортным сетям.

Положительный эффект также тесно связан с местоположением фирмы,

независимо от ее принадлежности к ОЭЗ, по мере удаления от экономического центра субъекта, падения концентрации спроса и предложения фирма теряет доступ к таким преимуществам. Показана связь между двумя данными эффектами. Однако остается открытым вопрос о величине такого взаимодействия, пока можно утверждать, что синергетический эффект, достигаемый за счет вступления фирмы в ОЭЗ и близости расположения к городским центрам экономической активности, остается положительным.

Вторая гипотеза заключалась в предположении относительно модели генерации резидентов, которая была получена благодаря применению метода *HistGBC*. Результат тестирования гипотезы не выявил однозначной взаимосвязи между местоположением ОЭЗ и решением фирмы о вступлении или невступлении в нее.

Таким образом, вторая гипотеза не может быть ни принята, ни отвергнута. Требуется дополнительное исследование в этой области.

Полученные результаты соотносятся с результатами Chaudey et al. [13], где была установлена положительная взаимосвязь между степенью близости организации к центру экономической активности и формируемым экономическим результатом (авторы оценивали влияние на переменную «*Economic growth*» и «*Economic return*»). Более того, в работе выявлены статистически значимые сетевые эффекты, оказывающие положительное влияние на деятельность организаций – участниц сетевого взаимодействия. Комбинация двух данных эффектов позволяет достичь синергетического результата. Однако авторы [13] отмечают, что издержки вступления и поддержания сети могут оказать губительный эффект на индивидуальном уровне для показателя выручки, не принося компании дополнительных доходов в сравнении с независимыми участниками рынка, стоит отметить, что такой

результат может быть обусловлен смещением, возникающим из-за большего количества наблюдений для фирм, расположенных на удаленных территориях.

В работе были также выявлены пространственные эффекты – зависимость среднегодового темпа прироста резидентов и частных инвестиций от региона расположения ОЭЗ. Среди исследователей и государственных деятелей существует мнение, что региональным диспропорциям присущ механизм положительной обратной связи. Эффект может быть подтвержден, однако сила влияния умеренная.

К ограничениям работы можно отнести тот факт, что наши результаты специализируются главным образом на особых экономических зонах промышленно-производственного и портового типов, что не позволяет их в полной мере распространить на всю совокупность особых зон, расположенных в пределах Российской Федерации. Требуется проведение дополнительного сбора данных и расширения сформированной выборки организаций – резидентов ОЭЗ, что может являться предметом будущих исследований по данной теме.

6. Заключение

В исследовании была подтверждена гипотеза о наличии положительной связи между местоположением организации и ее формируемым по итогам года финансовым результатом. Резиденты ОЭЗ в России приобретают конкурентные преимущества благодаря предоставленным им преференциям в сравнении с независимыми участниками рынка. В среднем по выборке компании-резиденты, расположенные вблизи центров субъектов РФ, более успешны по показателю генерируемой выручки.

Ведение предпринимательской деятельности компанией на территории ОЭЗ имеет значение: выборки предприятий-резидентов и нерезидентов действительно различны

согласно непараметрическому критерию Уилкоксона.

Гипотеза о наличии сильной взаимосвязи между показателями пространственной дифференциации субъектов РФ и показателями, характеризующими процесс генерации резидентов и частных инвестиций, не была в полной мере подтверждена, наблюдаемая связь умеренно положительная.

Процесс генерации резидентов может быть успешно объяснен рядом количественных регрессоров. Положительное влияние оказывают близость расположения к центрам субъектов РФ, объем налоговых льгот, виды экономической деятельности резидентов, тогда как отрицательно влияют протяженность автомобильных дорог и объем уплаченных налогов.

Предсказать вероятность вхождения организации в состав ОЭЗ представляется возможным благодаря применению логистической регрессии и классификационных алгоритмов машинного обучения, для работы с данными, имеющими пропуски значений переменных. Более всего подходит алгоритм *Histogram-based Gradient Boosting Classification Tree*.

Выявлен ряд количественных факторов, с помощью которых возможно составить карту предсказаний для организаций – потенциальных резидентов. Среди них местоположение и тип ОЭЗ, объем льгот, предоставляемых

резидентам, и налогов, мощность объектов теплоснабжения, стоимость электроснабжения на территории зоны, коэффициенты автономии, рентабельности активов, собственного капитала, продаж (которые могут быть представлены в бизнес-плане компании).

Теоретическая значимость исследования заключается в спецификации и тестировании моделей, обеспечивающих возможность анализа процесса генерации резидентов и принятия решений компанией, потенциальным резидентом ОЭЗ.

Практическая значимость состоит в возможности использования полученных результатов при разработке механизмов стимулирования инвестиционных проектов, результаты исследования могут быть применены на практике управляющими компаниями ОЭЗ при привлечении потенциальных резидентов и предпочтений на результаты формируемого оборота по итогам финансового года.

Предложенные в работе подходы могут быть полезны потенциальным резидентам, инвесторам и управляющим компаниям особых экономических зон. Описанный поэтапно процесс моделирования может быть применен организациями для решения вопроса о вхождении в состав ОЭЗ, управляющими компаниями – для моделирования эффектов, влияющих на процесс генерации резидентов и частных инвестиций.

Список использованных источников

1. Ramos-Rodríguez A., Medina-Garrido J., Lorenzo-Gómez J., Ruiz-Navarro J. What you know or who you know? The role of intellectual and social capital in opportunity recognition // International Small Business Journal. 2010. Vol. 28, Issue 6. Pp. 566–582. <https://doi.org/10.1177/0266242610369753>
2. Радыгина С. В., Суворова В. В. Особые экономические зоны как инструмент экономического роста и развития промышленного потенциала России // Вестник Удмуртского университета. Серия «Экономика и право». 2020. Т. 30, №. 3. С. 346–353. <https://doi.org/10.35634/2412-9593-2020-30-3-346-353>
3. Иванов С. А., Кожевникова П. А. Особые экономические зоны Дальнего Востока России: децентрализованная и централизованная модели управления // Труды института истории, археологии и этнографии ДВО РАН. 2019. Т. 24, № 3. С. 161–176. <https://doi.org/10.24411/2658-5960-2019-10034>

4. *Квашнина И. А.* Особые экономические зоны как инструмент привлечения иностранных инвестиций // Вестник Института экономики Российской академии наук. 2021. №. 3. С. 121–132. https://doi.org/10.52180/2073-6487_2021_3_121_132
5. *Ахмедзянов Р. Р., Алексеева Е. В., Филиппов М. И., Мосеевкова А. О.* Таможенно-тарифное регулирование в особых экономических зонах // Московский экономический журнал. 2019. №. 13. С. 89–96. <https://doi.org/10.24411/2413-046X-2019-10282>
6. *Какаулина М. О., Горлов Д. Р.* Оценка влияния налоговых льгот на инвестиционную активность в особых экономических зонах Российской Федерации // Journal of Applied Economic Research. 2022. Т. 21, № 2. С. 282–324. <https://doi.org/10.15826/vestnik.2022.21.2.011>
7. *Min J., Kang B.* Promoting New Growth: ‘Advanced Special Economic Zones’ in the Russian Far East // In: *Russia’s Turn to the East. Global Reordering.* Edited by H. Blakkisrud, E. Wilson Rowe. Palgrave Pivot, 2018. Pp. 51–74. https://doi.org/10.1007/978-3-319-69790-1_4
8. Бизнес-навигатор по особым экономическим зонам России – 2022. Выпуск 6 / Д. А. Гуляева, И. Е. Гусев, Е. А. Баскакова, М. М. Бухарова, Е. А. Князева, Е. И. Кравченко, М. А. Лабудин, М. К. Мальбахов, Е. А. Парамзина, А. П. Семенов, А. В. Шпиленко. М.: АКИТ РФ, 2022. 251 с. URL: https://www.economy.gov.ru/material/file/699ec37679f67c137b-011926f7a15119/Business_Navigator_2022.pdf
9. Бизнес-навигатор по особым экономическим зонам России – 2021. Выпуск 5 / М. М. Бухарова, А. Н. Андреев, Р. Ф. Бододько, Д. А. Гуляева, В. И. Зверков, Е. И. Кравченко, М. А. Лабудин, М. К. Мальбахов, А. Р. Новикова, М. С. Серегин, В. А. Суров, А. В. Шпиленко. М.: АКИТ РФ, 2021. 265 с. URL: https://economy.gov.ru/material/file/069f5d-d4923b44519322fc84cd9ebcfb/Business_Navigator_2021.pdf
10. *Корева О. В., Тухий В. И., Тяпкина А. П.* Исследование особых экономических зон в контексте устойчивого развития регионов // Дискуссия. 2022. Т. 113, №. 4. С. 52–66. <https://doi.org/10.46320/2077-7639-2022-4-113-52-65>
11. *Кузнецова О. В.* Научно-технологические приоритеты в федеральной политике пространственного развития в России // Федерализм. 2023. Т. 27, №. 4. С. 5–20. <https://doi.org/10.21686/2073-1051-2022-4-5-20>
12. *Доржиева В. В., Сорокина Н. Ю., Беляевская-Плотник Л. А., Волкова Н. Н., Романюк Э. И.* Пространственные аспекты инновационного и научно-технологического развития России: Научный доклад. М.: ИЭ РАН, 2022. 94 с. URL: https://inecon.org/docs/2022/Spatial_aspects_innovative_scientific-technological_development_Russia.pdf
13. *Chaudey M., Fadaïro M., Perdreau F.* Do retailers benefit from network affiliation in all locations // Economics Bulletin. 2020. Vol. 40, Issue 2. Pp. 1623–1633. URL: <http://www.accessecon.com/Pubs/EB/2020/Volume40/EB-20-V40-I2-P139.pdf>
14. *Неретина Е. А.* Типы, конфигурация и способы построения межорганизационных сетей // Известия высших учебных заведений. Поволжский регион. Общественные науки. 2014. №. 2 (30). С. 196–204. URL: <https://rucont.ru/efd/552313>
15. *Blois K. J.* Transaction costs and networks // Strategic Management Journal. 1990. Vol. 11, Issue 6. Pp. 493–496. <https://doi.org/10.1002/smj.4250110607>
16. *Saar-Tsechansky M., Provost F.* Handling missing values when applying classification models // Journal of Machine Learning Research. 2007. Vol. 8. Pp. 1625–1657. <https://doi.org/10.5555/1314498.1314553>
17. *Suthaharan S.* Support Vector Machine // In: *Machine Learning Models and Algorithms for Big Data Classification.* Integrated Series in Information Systems. Vol. 36. Boston, MA: Springer, 2016. Pp. 207–235. https://doi.org/10.1007/978-1-4899-7641-3_9
18. *Rokach L., Maimon O.* Decision Trees // In: *Data Mining and Knowledge Discovery Handbook.* Edited by O. Maimon, L. Rokach. Boston, MA: Springer, 2005. Pp. 165–192. https://doi.org/10.1007/0-387-25465-X_9
19. *Murphy K. P.* Naive bayes classifiers // University of British Columbia. 2006. Vol. 18, No. 60. Pp. 1–8. URL: <https://www.ic.unicamp.br/~rocha/teaching/2011s2/mc906/aulas/naive-bayes.pdf>

20. Mucherino A., Papajorgji P.J., Pardalos P.M. k-Nearest Neighbor Classification // In: Data Mining in Agriculture. Springer Optimization and Its Applications. Vol. 34. Pp. 83–106. New York, NY: Springer, 2009. https://doi.org/10.1007/978-0-387-88615-2_4
21. Peter S., Diego F., Hamprecht F.A., Nadler B. Cost efficient gradient boosting // Proceedings of the 31st International Conference on Neural Information Processing Systems (NIPS 2017). Long Beach, CA, 2017. URL: https://proceedings.neurips.cc/paper_files/paper/2017/file/4fac9ba115140ac4f1c22da82aa0bc7f-Paper.pdf
22. Chaudhary A., Kolhe S., Kamal R. An improved random forest classifier for multi-class classification // Information Processing in Agriculture. 2016. Vol. 3, Issue 4. Pp. 215–222. <https://doi.org/10.1016/j.inpa.2016.08.002>
23. Ke G., Meng Q., Finley T., Wang T., Chen W., Ma W., Ye Q., Liu T. LightGBM: A Highly Efficient Gradient Boosting Decision Tree // Proceedings of the 31st International Conference on Neural Information Processing Systems (NIPS 2017). Long Beach, CA, 2017. Pp. 3149–3157. URL: https://proceedings.neurips.cc/paper_files/paper/2017/file/6449f44a102fde848669bdd9e-b6b76fa-Paper.pdf
24. Tamim Kashifi M., Ahmad I. Efficient histogram-based gradient boosting approach for accident severity prediction with multisource data // Transportation Research Record. 2022. Vol. 2676, Issue 6. Pp. 236–258. <https://doi.org/10.1177/03611981221074370>
25. Marlin B.M. Missing Data Problems in Machine Learning. University of Toronto, 2008. URL: https://people.cs.umass.edu/~marlin/research/phd_thesis/marlin-phd-thesis.pdf
26. Guryanov A. Histogram-based algorithm for building gradient boosting ensembles of piecewise linear decision trees // Proceedings of 8th International Conference on Analysis of Images, Social Networks and Texts. AIST 2019. Springer, 2019. Pp. 39–50. https://doi.org/10.1007/978-3-030-37334-4_4
27. Mansfield E.R., Helms B.P. Detecting multicollinearity // The American Statistician. 1982. Vol. 36, Issue 3a. Pp. 158–160. <https://doi.org/10.1080/00031305.1982.10482818>
28. Kleiber C., Zeileis A. Applied Econometrics with R. New York, NY: Springer, 2015. <http://dx.doi.org/10.1007/978-0-387-77318-6>
29. Deepa B., Ramesh K. Epileptic seizure detection using deep learning through min max scaler normalization // International Journal of Health Sciences. 2022. Vol. 6. Pp. 10981–10996. <https://doi.org/10.53730/ijhs.v6nS1.7801>
30. Fay M.P., Proschan M.A. Wilcoxon-Mann-Whitney or t-test? On assumptions for hypothesis tests and multiple interpretations of decision rules // Statistics Surveys. 2010. Vol. 4. Pp. 1–39. <https://doi.org/10.1214/09-SS051>
31. Ng H.K.T., Balakrishnan N., Panchapakesan S. Selecting the Best Population Using a Test for Equality Based on Minimal Wilcoxon Rank-sum Precedence Statistic // Methodology and Computing in Applied Probability. 2007. Vol. 9. Pp. 263–305. <https://doi.org/10.1007/s11009-007-9023-9>
32. Дебердиева Е.М., Вечкасова М.В., Фролова С.В. Совершенствование регулирования производств высокой добавленной стоимости // Вестник Южно-Уральского государственного университета. Серия: Экономика и менеджмент. 2021. Т. 15, №. 3. С. 56–63. <https://doi.org/10.14529/em210306>
33. Benini R., Czyzewski A. Regional disparities and economic growth in Russia: new growth patterns and catching up // Economic Change and Restructuring Volume. 2007. Vol. 40. Pp. 91–135. <https://doi.org/10.1007/s10644-007-9026-0>
34. Yershov Yu.S. Features of regional economic development in Russia in 1999–2013 // Regional Research of Russia. 2016. Vol. 6, Issue 4. Pp. 281–291. <https://doi.org/10.1134/S2079970516040079>
35. Schober P., Boer C., Schwarte L.A. Correlation coefficients: appropriate use and interpretation // Anesthesia & Analgesia. 2018. Vol. 126, Issue 5. Pp. 1763–1768. <https://doi.org/10.1213/ane.0000000000002864>

ИНФОРМАЦИЯ ОБ АВТОРАХ

Плесовских Александр Евгеньевич

Лаборант-исследователь лаборатории экономики климатических изменений и экологического развития Сибирского федерального университета, г. Красноярск, Россия (660041, г. Красноярск, пр. Свободный, 79); ORCID <https://orcid.org/0000-0001-8507-9501> e-mail: alexandermcme@gmail.com

ДЛЯ ЦИТИРОВАНИЯ

Плесовских А.Е. Особые экономические зоны Российской Федерации: моделирование решений потенциальных резидентов и процесса их генерации // Journal of Applied Economic Research. 2023. Т. 22, № 2. С. 323–354. <https://doi.org/10.15826/vestnik.2023.22.2.014>

ИНФОРМАЦИЯ О СТАТЬЕ

Дата поступления 31 января 2023 г.; дата поступления после рецензирования 23 февраля 2023 г.; дата принятия к печати 29 марта 2023 г.

Special Economic Zones of Russia: Forecasting Decisions of Potential Residents and Resident Generation Process Modeling

Alexander E. Plesovskikh  

Siberian Federal University,
Krasnoyarsk, Russia

 alexandermcme@gmail.com

Abstract. Modern studies widely discuss the role of special economic zones in stimulating the economic growth and development of Russia, generating the necessary investment flows and increasing the country's innovative potential by expanding production in high-tech sectors of the economy with high added value. The purpose of the study is to model the process of generating residents and to determine quantitative factors that have a statistically significant effect on the average annual growth rate of companies participating in special economic zones in the Russian Federation. The paper describes modern approaches to predicting the choice of potential residents to start doing business in the territory of the SEZ using classification approaches (Support Vector Machines, Decision Trees, Random Forest, Naive Bayes, K-Nearest Neighbor, Gradient Boosting) and regression approaches (logistic regression). A modern classification algorithm was applied in practice - Histogram-based Gradient Boosting Classification Tree, which is stable for analyzing large data with missing variable values and does not require preliminary sample transformation. The paper confirms the hypothesis that there is a positive relationship between the location of the organization and its financial result forming by the end of the year. On average, in the sample, resident companies located near the centers of the constituent entities of the Russian Federation are more successful in terms of generated revenue. The hypothesis that there is a strong relationship between indicators of spatial differentiation of the regions of the Russian Federation and indicators characterizing the process of generating residents and private investment has not been fully confirmed. From a practical point of view, the results of the study could be applied by both resident organizations, potential residents, and SEZ management companies. The theoretical significance of the study lies in the specification of the proposed binary choice model for potential residents, which can be expanded and generalized in future works. At present, there are all the necessary prerequisites for creating conditions for the development of industry, high-tech sectors of the economy and the production of high value-added products in order to increase the stability of the Russian economy.

Key words: Russian special economic zones; resident generation; machine learning; regression and classification; binary choice models.

JEL C12, C23, O11, O47

References

1. Ramos-Rodríguez, A., Medina-Garrido, J., Lorenzo-Gómez, J., Ruiz-Navarro, J. (2010). What you know or who you know? The role of intellectual and social capital in opportunity recognition. *International Small Business Journal*, Vol. 28, Issue 6, 566–582. <https://doi.org/10.1177/0266242610369753>
2. Radygina, S.V., Suvorova, V.V. (2020). Osoby ekonomicheskie zony kak instrument ekonomicheskogo rosta i razvitiia promyshlennogo potentsiala Rossii (Special economic zones as an instrument of economic growth and development of industrial potential of Russia). *Vestnik Udmurtskogo universiteta. Seriya «Ekonomika i pravo» (Bulletin of Udmurt University. Series*

Economics and Law), Vol. 30, No. 3, 346–353. (In Russ.). <https://doi.org/10.35634/2412-9593-2020-30-3-346-353>

3. Ivanov, S.A., Kozhevnikova, P.A. (2019). Osobyie ekonomicheskie zony Dal'nego Vostoka Rossii: detsentralizovannaia i tsentralizovannaia modeli upravleniia (Special economic zones of the Russian Far East: Decentralized and centralized models). *Trudy instituta istorii, arkhologii i etnografii DVO RAN (Proceedings of the Institute of History, Archaeology and Ethnology FEB RAS)*. Vol. 24, No. 3, 161–176. (In Russ.). <https://doi.org/10.24411/2658-5960-2019-10034>

4. Kvashnina, I.A. (2021). Osobyie ekonomicheskie zony kak instrument privlecheniia inostrannykh investitsii (Special economic zones as a tool to attract foreign investments). *Vestnik Instituta ekonomiki Rossiiskoi akademii nauk (Bulletin of the IE RAS)*, No. 3, 121–132. (In Russ.). https://doi.org/10.52180/2073-6487_2021_3_121_132

5. Akhmedzianov, R.R., Alekseeva, E.V., Filippov, M.I., Moseenkova, A.O. (2019). Tamozhenno-tarifnoe regulirovanie v osobykh ekonomicheskikh zonakh (Customs and tariff regulation in special economic zones). *Moskovskii ekonomicheskii zhurnal (Moscow Economic Journal)*, No. 13, 89–96. (In Russ.). <https://doi.org/10.24411/2413-046X-2019-10282>

6. Kakaulina, M.O., Gorlov, D.R. (2022). Assessment of the Impact of Tax Incentives on Investment Activity in Special Economic Zones of the Russian Federation. *Journal of Applied Economic Research*, Vol. 21, No. 2, 282–324. (In Russ.). <https://doi.org/10.15826/vestnik.2022.21.2.011>

7. Min, J., Kang, B. (2018). Promoting New Growth: 'Advanced Special Economic Zones' in the Russian Far East. In: *Russia's Turn to the East. Global Reordering*. Edited by H. Blakkisrud, E. Wilson Rowe. Palgrave Pivot, 51–74. https://doi.org/10.1007/978-3-319-69790-1_4

8. Gulyayeva, I.E. et al. (2022). *Biznes-navigator po osobym ekonomicheskim zonam Rossii – 2022 [Business Navigator for Russia's special economic zones 2022]*. Issue 6. Moscow, AKIT RF. (In Russ.). Available at: https://www.economy.gov.ru/material/file/699ec37679f67c137b-011926f7a15119/Business_Navigator_2022.pdf

9. Bukharova, M.M. et al. (2022). *Biznes-navigator po osobym ekonomicheskim zonam Rossii – 2021 [Business Navigator for Russia's special economic zones 2021]*. Issue 5. Moscow, AKIT RF. (In Russ.). Available at: https://economy.gov.ru/material/file/069f5dd4923b44519322f-c84cd9ebcfb/Business_Navigator_2021.pdf

10. Koreva, O.V., Tikhii, V.I., Tiapkina, A.P. (2022). Issledovanie osobykh ekonomicheskikh zon v kontekste ustoichivogo razvitiia regionov (Research of special economic zones in the context of sustainable development of regions). *Diskussiiia (Discussion)*, Vol. 113, No. 4, 52–66. (In Russ.). <https://doi.org/10.46320/2077-7639-2022-4-113-52-65>

11. Kuznetsova, O.V. (2023). Nauchno-tehnologicheskie prioritety v federalnoi politike prostranstvennogo razvitiia v Rossii (Scientific and Technological Priorities in the Federal Spatial Development Policy in Russia). *Federalizm (Federalism)*, Vol. 27, No. 4, 5–20. (In Russ.). <https://doi.org/10.21686/2073-1051-2022-4-5-20>

12. Dorzhieva, V.V., Sorokina, N.Iu., Beliaevskaia-Plotnik, L.A., Volkova, N.N., Romaniuk, E.I. (2022). *Prostranstvennye aspekty innovatsionnogo i nauchno-tehnologicheskogo razvitiia Rossii (Spatial aspects of innovative and scientific-technological development of Russia)*. Moscow, Institute of Economics of Russian Academy of Sciences. (In Russ.). Available at: https://inecon.org/docs/2022/Spatial_aspects_innovative_scientific-technological_development_Russia.pdf

13. Chaudey, M., Fadairo, M., Perdreau, F. (2020). Do retailers benefit from network affiliation in all locations. *Economics Bulletin*, Vol. 40, Issue 2, 1623–1633. Available at: <http://www.accessecon.com/Pubs/EB/2020/Volume40/EB-20-V40-I2-P139.pdf>

14. Neretina, E.A. (2014). Tipy, konfiguratsiia i sposoby postroeniia mezhhorganizatsionnykh setei (Types, configuration and methods of building interorganizational networks). *Izvestiia vysshikh uchebnykh zavedenii. Povolzhskii region. Obschestvennye nauki (University pro-*

ceedings. *Volga region. Social sciences*, No. 2 (30), 196–204. (In Russ.). Available at: <https://ru-cont.ru/efd/552313>

15. Blois, K.J. (1990). Transaction costs and networks. *Strategic Management Journal*, Vol. 11, Issue 6, 493–496. <https://doi.org/10.1002/smj.4250110607>

16. Saar-Tsechansky, M., Provost, F. (2007). Handling missing values when applying classification models. *Journal of Machine Learning Research*, Vol. 8, 1625–1657. <https://doi.org/10.5555/1314498.1314553>

17. Suthaharan, S. (2016). Support Vector Machine. In: *Machine Learning Models and Algorithms for Big Data Classification. Integrated Series in Information Systems*, Vol. 36. Boston, MA, Springer, 207–235. https://doi.org/10.1007/978-1-4899-7641-3_9

18. Rokach, L., Maimon, O. (2005). Decision Trees. In: *Data Mining and Knowledge Discovery Handbook*. Edited by O. Maimon, L. Rokach. Boston, MA, Springer, 165–192. https://doi.org/10.1007/0-387-25465-X_9

19. Murphy, K.P. (2006). Naive bayes classifiers. *University of British Columbia*, Vol. 18, No. 60, 1–8. Available at: <https://www.ic.unicamp.br/~rocha/teaching/2011s2/mc906/aulas/naive-bayes.pdf>

20. Mucherino, A., Papajorgji, P.J., Pardalos, P.M. (2009). k-Nearest Neighbor Classification. In: *Data Mining in Agriculture. Springer Optimization and Its Applications*, Vol. 34, 83–106. New York, NY, Springer. https://doi.org/10.1007/978-0-387-88615-2_4

21. Peter, S., Diego, F., Hamprecht, F.A., Nadler, B. (2017). Cost efficient gradient boosting. *Proceedings of the 31st International Conference on Neural Information Processing Systems (NIPS 2017)*. Long Beach, CA. Available at: https://proceedings.neurips.cc/paper_files/paper/2017/file/4fac9ba115140ac4f1c22da82aa0bc7f-Paper.pdf

22. Chaudhary, A., Kolhe, S., Kamal, R. (2016). An improved random forest classifier for multi-class classification. *Information Processing in Agriculture*, Vol. 3, Issue 4. Pp. 215–222. <https://doi.org/10.1016/j.inpa.2016.08.002>

23. Ke, G., Meng, Q., Finley, T., Wang, T., Chen, W., Ma, W., Ye, Q., Liu, T. (2017). LightGBM: A Highly Efficient Gradient Boosting Decision Tree. *Proceedings of the 31st International Conference on Neural Information Processing Systems (NIPS 2017)*. Long Beach, CA, 3149–3157. Available at: https://proceedings.neurips.cc/paper_files/paper/2017/file/6449f44a102fde848669bdd9eb6b76fa-Paper.pdf

24. Tamim Kashifi, M., Ahmad, I. (2022). Efficient histogram-based gradient boosting approach for accident severity prediction with multisource data. *Transportation Research Record*, Vol. 2676, Issue 6, 236–258. <https://doi.org/10.1177/03611981221074370>

25. Marlin, B.M. (2008). *Missing Data Problems in Machine Learning*. University of Toronto. Available at: https://people.cs.umass.edu/~marlin/research/phd_thesis/marlin-phd-thesis.pdf

26. Guryanov, A. (2019). Histogram-based algorithm for building gradient boosting ensembles of piecewise linear decision trees. *Proceedings of 8th International Conference on Analysis of Images, Social Networks and Texts. AIST 2019*. Springer, 39–50. https://doi.org/10.1007/978-3-030-37334-4_4

27. Mansfield, E.R., Helms, B.P. (1982). Detecting multicollinearity. *The American Statistician*, Vol. 36, Issue 3a, 158–160. <https://doi.org/10.1080/00031305.1982.10482818>

28. Kleiber, C., Zeileis, A. (2015). *Applied Econometrics with R*. New York, NY, Springer. <http://dx.doi.org/10.1007/978-0-387-77318-6>

29. Deepa, B., Ramesh, K. (2022). Epileptic seizure detection using deep learning through min max scaler normalization. *International Journal of Health Sciences*, Vol. 6, 10981–10996. <https://doi.org/10.53730/ijhs.v6nS1.7801>

30. Fay, M.P., Proschan, M.A. (2010). Wilcoxon-Mann-Whitney or t-test? On assumptions for hypothesis tests and multiple interpretations of decision rules. *Statistics Surveys*, Vol. 4, 1–39. <https://doi.org/10.1214/09-SS051>

31. Ng, H.K.T., Balakrishnan, N., Panchapakesan, S. (2007). Selecting the Best Population Using a Test for Equality Based on Minimal Wilcoxon Rank-sum Precedence Statistic. *Methodology and Computing in Applied Probability*, Vol. 9, 263–305. <https://doi.org/10.1007/s11009-007-9023-9>

32. Deberdieva, E.M., Vechkasova, M.V., Frolova, S.V. (2021). Sovershenstvovanie regulirovaniia proizvodstv vysokoi dobavlennoi stoimosti (Improving the regulation of high value added industries). *Vestnik Iuzhno-Uralskogo gosudarstvennogo universiteta. Seriya: Ekonomika i menedzhment (Bulletin of South Ural State University, Series "Economics and Management")*, Vol. 15, No. 3, 56–63. (In Russ.). <https://doi.org/10.14529/em210306>

33. Benini, R., Czyzewski, A. (2007). Regional disparities and economic growth in Russia: new growth patterns and catching up. *Economic Change and Restructuring Volume*, Vol. 40, 91–135. <https://doi.org/10.1007/s10644-007-9026-0>

34. Yershov, Yu.S. (2016). Features of regional economic development in Russia in 1999–2013. *Regional Research of Russia*, Vol. 6, Issue 4, 281–291. <https://doi.org/10.1134/S2079970516040079>

35. Schober, P., Boer, C., Schwarte, L.A. (2018). Correlation coefficients: appropriate use and interpretation. *Anesthesia & Analgesia*, Vol. 126, Issue 5, 1763–1768. <https://doi.org/10.1213/ane.0000000000002864>

INFORMATION ABOUT AUTHOR

Alexander Evgenievich Plesovskikh

Research Assistant, Laboratory for Economics of Climate Change and Environmental Development, Siberian Federal University, Krasnoyarsk, Russia (660041, Krasnoyarsk, Svobodny Avenue, 79); ORCID <https://orcid.org/0000-0001-8507-9501> e-mail: alexandermeme@gmail.com

FOR CITATION

Plesovskikh, A.E. (2023). Special Economic Zones of Russia: Forecasting Decisions of Potential Residents and Resident Generation Process Modeling. *Journal of Applied Economic Research*, Vol. 22, No. 2, 323–354. <https://doi.org/10.15826/vestnik.2023.22.2.014>

ARTICLE INFO

Received January 31, 2023; Revised February 23, 2023; Accepted March 29, 2023.

